BULLETIN of the
SCHOOL OF EDUCATION
INDIANA UNIVERSITY

Skill in Teaching, A Reappraisal of the Concepts and Strategies in Teacher Effectiveness Research

by
RICHARD L. TURNER
and
NICHOLAS A. FATTU

Vol. 36, No. 3 May, 1960

OFFICERS OF THE UNIVERSITY

HERMAN B WELLS, A.M., LL.D., President
RALPH L. COLLINS, Ph.D., Vice-President, and Dean of the Faculties
JOSEPH AMOS FRANKLIN, B.S., Vice-President, and Treasurer
JOHN WILLIAM ASHTON, Ph.D., LL.D., Vice-President for Graduate Development, and Dean of the Graduate School
SAMUEL EDWARD BRADEN, Ph.D., Vice-President, and Dean for Undergraduate Development
CHARLES EDWIN HARRELL, A.B., LL.B., Registrar, and Director of the Office of Records and Admissions

ADMINISTRATIVE STAFF OF THE SCHOOL OF EDUCATION

HAROLD GRAY SHANE, Ph.D., Dean
HOWARD T. BATCHELDER, Ph.D., Associate Dean
PHILIP PEAK, Ph.D., Assistant Dean
(MRS.) DOROTHY J. MCCAW, Assistant to the Dean
OTTO HUGHES, Ed.D., Director of the University School
KATHLEEN DUGDALE, A.M., Editor of Publications in Education

ADMINISTRATIVE STAFF OF THE DIVISION OF RESEARCH AND FIELD SERVICES

HAROLD H. CHURCH, Ph.D., Director
H. GLENN LUDLOW, Ph.D., Assistant Director
RUTH FAY ARGANBRIGHT, M.S. in Ed., Assistant Director
W. MONFORT BARR, Ed.D., Assistant Director of School Surveys
PAUL W. SEAGERS, Ed.D., School Building Consultant

BULLETIN OF THE SCHOOL OF EDUCATION
INDIANA UNIVERSITY

SKILL IN TEACHING, A REAPPRAISAL OF THE CONCEPTS AND STRATEGIES IN TEACHER EFFECTIVENESS RESEARCH

by

Richard L. Turner

and

Nicholas A. Fattu

Published by the

DIVISION OF RESEARCH AND FIELD SERVICES

Indiana University

Price: $1.25 per copy postpaid
Available at the Indiana University Bookstore
Bloomington, Indiana

Subscriptions
$5.00 per year, for six issues
Subscribe through the School of Education,
Indiana University, Bloomington, Indiana

Entered as second-class mail matter September 30, 1924, at the post office at Bloomington, Indiana, under the Act of August 24, 1912. Published in January, March, May, July, September, and November of each year, from the University Office, Bloomington, Indiana.

Preface

About ten years ago, the staff of the Institute of Educational Research, in cooperation with the elementary and secondary supervisors at Indiana University, attempted an evaluation of student teaching outcomes. Results of more than three years of study were highly specific and relevant only to the situation in which they were collected. Virtually no generalizations were noted that could be used as dependable guides in dealing with parallel situations. The high specificity of these findings was attributed to unsatisfactory identification and definition (hence control) of the variables involved.

In an attempt to define variables capable of yielding more generalizable results, a comprehensive survey of the published literature was made. Although this literature was enormous, the results reported were generally as specific as the Institute's studies. Most of the studies dealt, on a global basis, with the tremendously complex interactions found in actual classrooms. Attempts to identify, define, and then rigorously control the variables seemed inevitably to lead to the same frustration encountered by the Institute. Seventy years of research on teacher effectiveness have not added much to our systematic knowledge, and it is difficult to see how another seventy can do any more if the same procedures are followed. When every study must virtually start anew at the same place as its predecessor, little increase can be made in organized knowledge in the field.

Parallels of this state of affairs can be found in every field. For example, in medicine a quarter of a century ago the prognosis for a patient with pneumonia was very poor. Practicing physicians applied their existing knowledge and remedies with relatively little success. Then, as a result of inquiry into very limited phenomena conducted under very restricted laboratory conditions, antibiotics were discovered, and the outlook for pneumonia patients changed completely. Similarly, in agriculture, study and revision of cultural practices increased the yield of corn somewhat. However, the development of hybrid corn by plant geneticists, working under rigidly controlled conditions with very restricted phenomena, greatly improved yield expectations. These are typical rather than isolated examples. The history of science is a series of examples such as these. Advancement in knowledge comes typically from breaking things down into simple components that can be studied rigorously, and not from studying organized wholes.

The methods of professional practitioners are not necessarily congruent with those of scientists. Practitioners must deal with the entire complex of entire situations. Their function is not to make advances in knowledge, but to apply current

knowledge to the best of their ability to the performance of the essential public service that is their specialty.

Educational research cannot effectively produce new knowledge if it continues to emulate the methods of the professional practitioner. To provide cumulative systematic knowledge, it must deal with carefully defined restricted phenomena studied under rigorously controlled conditions.

The research reports, of which the present is an example, are intended primarily to illustrate how an aspect of teaching performance can be more adequately studied than heretofore. These reports are designed to illustrate a procedure and to present results that can be combined with many other results obtained under parallel conditions to develop a systematic body of knowledge about teacher problem-solving behavior. The greatest value of these reports will be found only if they stimulate other educational researchers to develop alternative methods, procedures, and data.

R.L.T.
N.A.F.

TABLE OF CONTENTS

Page

Introduction 1

QUESTION 1: How Well Does a Particular Elementary School Teacher Perform His Particular Job? . . . 2

QUESTION 2: How Can One Predict How Well a Particular Elementary School Teacher Will Perform a Particular Job in Elementary School Teaching? . 7

QUESTION 3: How Well Does a Particular Elementary School Teacher Solve Those Problems Which Characterize Elementary School Teaching as a Professional Endeavor But Which Do Not Equally Well Represent All Jobs in Elementary School Teaching? 13

QUESTION 4: How Can One Predict How Well a Particular Elementary School Teacher Will Solve Those Problems Which Characterize Elementary School Teaching as a Professional Endeavor But Which Do Not Equally Well Represent All Jobs in Elementary School Teaching? 32

The research on which this study is based was supported in part by the Cooperative Research Program of the U.S. Office of Education under Project 419 (7790).

Vita

Nicholas A. Fattu

Dr. Fattu received his Ph.D. degree from the University of Minnesota. He has held positions as mathematics teacher, Itasca Junior College; Research Statistician, Committee on Institutional Research, University of Minnesota; Psychologist, College Entrance Examination Board and Office of Scientific Research and Development; Associate Professor of Psychology, Michigan State College; and Chief, Job and Proficiency Analysis Divisions, Human Resources Research Center, U.S. Air Force. He is at present Professor of Educational Psychology and Director of the Institute of Educational Research at Indiana University. He is currently a member of the Research Planning Committee, Council for Research in Education.

Richard L. Turner

Dr. Turner received the B.S. and A.M. degrees from Northwestern University, and in 1959 received the Ph.D. degree from Indiana University. He has been a supervising teacher in the laboratory schools at Wisconsin State College at Whitewater, and at Antioch College, Yellow Springs, Ohio. He has been associated with the Institute of Educational Research at Indiana University since 1956.

SKILL IN TEACHING, A REAPPRAISAL OF THE CONCEPTS AND STRATEGIES IN TEACHER EFFECTIVENESS RESEARCH

Introduction

A focal problem in any research endeavor, whether in education or in the social or physical sciences, is to ask "good" questions. A good question in research serves a number of functions. First, it leads to clarification of the objectives toward which the research is directed. Second, it leads to answers which are verifiable and which are significant either to a body of theory or to practical affairs or to both. Third, it leads to clarification of the assumptions upon which the researcher bases his work. Fourth, it helps define the research strategies and the specific procedures which the researcher can adopt in his attempt to answer the questions. And fifth, it leads to additional questions, the answers to which extend the scope of systematic, organized knowledge.

The purpose of this study is to consider and appraise four related questions bearing on the skill an elementary school teacher* displays in discharging his professional functions in order (a) to disentangle the questions which have been confounded in research in teacher effectiveness, (b) to identify the questions toward which an important segment of research in teacher effectiveness is now directed, and (c) to point out a way to answer some of the questions which current research in teacher effectiveness does not answer.

The four questions to be considered are as follows:

1. How well does a particular elementary school teacher perform his particular job?

2. How can one predict how well a particular elementary school teacher will perform a particular job in elementary school teaching?

3. How well does a particular elementary school teacher solve those problems which characterize elementary school teaching as a professional endeavor but which do not equally well represent all jobs in elementary school teaching?

4. How can one predict how well a particular elementary school teacher will solve those problems which characterize elementary school teaching as a professional endeavor but which do not equally well represent all jobs in elementary school teaching?

---

*The term "teacher" in the present paper means elementary school teacher unless otherwise specified.

QUESTION 1: How Well Does a Particular Elementary School Teacher Perform His Particular Job?

In examining the question as to how well a particular elementary school teacher performs his particular job, it seems desirable to take two perspectives, the one that a superintendent or other administrative official might take and the one that a researcher in teacher effectiveness might take. The reason for examining the question in these two perspectives is that the two have not been made sufficiently distinct and that the failure to make them distinct is one of the major sources of difficulty in research in teacher effectiveness.

The perspective of the administrator. The concern over how well a teacher performs his job arises regularly in the context of practical affairs. The superintendent, supervisor, and principal are confronted with practical decisions concerning whether or not to make an extended effort to help teachers who perform poorly and, ultimately, whether or not to retain the services of such teachers.

By virtue of local control of schools, administrative officials are concerned with making personnel decisions about a limited number of teachers within a defined geographical area. They are not concerned with making decisions about the teacher personnel in some other city or some other county. Also by virtue of local control, the definition of the functions which a particular teacher is to discharge is left largely to one or more groups or combinations of groups in the community which supports the school system. These groups may include (a) administrative officials, (b) teachers, (c) the school board, and (d) persons of power and influence in the community.

Administrative officials cannot make a decision as to how well a particular teacher performs his job without defining, however loosely, what the job is. When the job is loosely defined, they may appeal to such criteria as how well they like the teacher personally, the extent to which the teacher disrupts or facilitates the smooth running of the administrative machinery, and the number of complaints or praises the teacher receives from parents.

At a somewhat more organized level, administrative officials may actually work with the various groups, including the school board, teachers, and community members, to determine what the functions of a teacher in the local school system should be. Such groups may rigidly specify function and provide little autonomy for the teacher, or they may state functions broadly and leave much teacher autonomy. Whatever the case, the group or groups are concerned with defining teacher function within a limited geographical area. Just as administrative officials make no personnel decisions for systems other than their own, such groups make no statements with regard to teacher function for other than their own system.

Local control of schools entails, then, that teacher function is free to vary from school system to school system, that is, the job of a teacher varies according to where the job is.

Now, even though the functions that a teacher is to discharge may be defined in a well organized way in a given school system, administrative officials must also consider the fact that, while first grade teachers, sixth grade teachers, and high school teachers may perform the same functions in a very general sense, their specific functions are quite different. Few school systems include among the functions of first grade teachers that of teaching quadradic equations, and few charge high school mathematics teachers with teaching a basic sight vocabulary in the English language. The particular job a teacher is to perform, then, varies according to grade level.

When function varies both with school systems and with grade levels, being a first grade teacher in a particular school system is by no means identical to being a fourth grade teacher in the same system, nor is it identical to being a first grade teacher in some other school system. In one school system the first grade teacher may be assigned a series of beginning books in reading and may be periodically checked by administrative officials who make certain his group is on the correct page in the correct book for that day in the semester. In another system, the teacher may be free to choose his own materials, his methods of using them, and his progress through them. The differences in the way specific teacher functions are defined in these two systems are very great. Yet, because an administrator is responsible primarily to a local school corporation, he need not be greatly concerned with the discrepancy. His decisions involve only the particular teachers performing the particular jobs in his system.

Because administrators are primarily responsible to a local school corporation and can assess the performance of a teacher by reference to teacher functions which are also locally defined, some of the procedures used for assessing teacher performance are not useful for other purposes. A full consideration of the extent to which these procedures are desirable for use by supervisors, principals, or superintendents is discussed elsewhere (8: ch. 11; 35: ch. 13)* and is not the primary concern of this paper.

Three general procedures for assessing teacher performance are generally recognized--ratings, systematic observation, and student gains (21: 3-4). Ratings may be sub-classified according to who does the rating. Administrative ratings may consist of an over-all rating of the teacher or of the rating of specific teacher behaviors or traits. An over-all rating may require that the administrator give the "per cent of effec-

---

*The references are listed by number in the bibliography on page 35.

tiveness" of the teacher (31: 2). More detailed ratings take a wide variety of forms, according to the purposes of the rating, and may be done by the teacher's peers or by students or administrative or staff personnel. Self-ratings by teachers may have much to commend them from the viewpoint of improving instruction (35: 298-303), although their use for other purposes may be questioned.

Systematic observations of the teacher may be used by administrative personnel, but the "practical" use of this technique probably conjuncts with the more systematic rating schedules insofar as administrators or supervisors are usually more interested in evaluating behavior than in describing it.

Finally, student gain, as measured by standardized tests, may be used to assess results purported to be attributable to what the teacher does in the classroom.

From the viewpoint of administrative officials, the extent to which any or all of the procedures which may be used in assessing the extent to which a particular teacher performs his particular job depends on how much and what kind of evidence the administrator desires in making his practical decisions with regard to local teacher personnel. His concerns do not extend beyond this point.

The perspective of the researcher in teacher effectiveness. The concern over how well a particular teacher performs his particular job arises for the educational researcher on the basis of one or the other or both of two considerations. The researcher either wishes to provide information which will help superintendents select teachers who will perform well or to provide information which will help teacher preparatory institutions select candidates who can be educated to perform well in subsequent teaching positions.

The most apparent strategy, and the recommended one (22) for gaining this information, consists, in broad outline, of first determining the individuals who perform well in teaching and then determining the behavioral characteristics that are associated with high performance. Thus, in the phase of the strategy which deals with how well teachers perform, it seems as if both researchers in teacher effectiveness and administrative officials in the schools have a common concern. Both are concerned with how well a particular teacher performs his particular job. However, a closer examination of the purposes of the researcher indicate that his concerns, while giving the appearance of similarity to those of the superintendent, are in fact quite different. The superintendent seeks to determine whether a particular teacher performs his job well; he is not concerned with whether the particular job he asks a teacher to perform is representative of a class of such jobs or whether the teacher can perform a class of jobs well. But the researcher is interested in whether a particular teacher can perform well in any of a class of jobs which share a great many common characteristics. Because his ultimate intent is to predict, the researcher must be concerned with the extent

to which the functions to be discharged in a given job represent those functions which are to be discharged in all the instances about which he wishes to make predictions.

The distinction between the administrator's concern and the researcher's concern has many implications. At this point it is focal only to examine how the differences in the concerns of the administrator and the researcher change the utility and relevance of the first mentioned procedure which may be used to assess teacher performance, the over-all administrative rating.

The over-all administrative rating may be used by an administrative official to aid in making a very general assessment of how well a teacher performs. It is usually considered relevant, although not necessarily desirable, for principals or supervisors to make a general, somewhat intuitive, appraisal of a teacher in relation to the functions he is to discharge. It is not equally relevant, however, for a researcher in teacher effectiveness to use such over-all ratings. This is the case not because such ratings have low reliabilities (21: 15), which make them undesirable though not irrelevant, but because over-all ratings are not consistent with the central purpose of the researcher: to predict. An over-all rating of effectiveness for research purposes implicitly assumes that there is no variation between jobs in teaching; that is, it assumes that when one superintendent asserts that a teacher is 90 per cent effective he means that the teacher is as effective in doing the same things as every other teacher who is rated 90 per cent effective by other superintendents. But an over-all scale does not show that the teachers are effective or ineffective in doing the same things, it only shows that they are effective or ineffective in doing whatever it is that they do. An over-all rating is simply a way of letting the criterion which the researcher wants to predict vary in its meaning, without showing that it so varies. To consistently predict a criterion the meaning of which changes constantly is an impossible and useless task.

At this point the researcher in teacher effectiveness is confronted with a paradox. If he uses an over-all rating scale, or any procedure which permits the definition of the teacher's function to vary within the criterion he wants to predict, he will not be successful in predicting because he lets the definition of the criterion vary. On the other hand, if he does not let the definition of the criterion vary, he places himself in the position of having to specify what the function of teachers should be. If he specifies what the functions of teachers should be, he either automatically usurps the function of local school corporations in deciding what functions a teacher should discharge, or else he runs the risk of predicting a criterion which some local school corporations consider inconsequential, irrelevant, or negatively related with respect to how they define teacher function.

An example of this paradox can be shown by reference to the manual for the Minnesota Teacher Attitude Inventory (MTAI) (11: 4), in which it is suggested that, if a school adminis-

trator who intends to use the MTAI for selection purposes scores low on this instrument, he may find that teachers who score high on the instrument do not fit well into the school system he administers. If a researcher chooses to define high MTAI scores as a criterion, he has chosen a criterion which, from the viewpoint of the school administrator who scores low on the MTAI, selects ineffective rather than effective teachers and thus holds a negative relationship to the administrator's criterion of effectiveness.

Basically there are two ways to resolve this paradox. Each involves different research procedures and is directed toward somewhat different purposes. To understand the differences in these approaches it is necessary to re-examine the reasons why researchers become concerned with how well a particular teacher performs his particular job in the first place.

As has been stated, the researcher in teacher effectiveness may be intent either upon providing information which will help superintendents select teachers who will perform well, or upon providing teacher preparatory institutions with information on the basis of which candidates can be selected. In the first case, the task of the researcher is essentially to match teachers with teaching jobs by virtue of having found out what behavioral characteristics of teachers are associated with performing well in a certain kind of teaching situation. By knowing what kinds of functions a given school system wants discharged and the conditions under which they want them discharged, the researcher could help the administrative officials select teachers whose characteristics are most closely associated with high performance of the functions that that school system wants discharged under the conditions that prevail there.

In the case of the researcher who wishes to supply useful information to teacher preparatory institutions, the problem is quite a different one. The institution places its graduates in a wide variety of school systems. It is not the function of such an institution to select and prepare teachers to succeed in only a particular school system, but rather to select and prepare teachers who, as professional persons, can solve teaching problems under a wide variety of conditions. It is thus the task of researchers who seek to provide information for preparatory institutions to identify those persons whose performance as teachers is not dependent upon their functions in a particular school system, but whose teaching skills permit successful performance in a variety of school situations. There are probably limits to the range of school situations in which a given teacher can succeed. To take a case in point, if a teacher scores high on the MTAI there are probably a great many situations in which he can succeed, other things being equal, but there are also some situations which deviate so far with respect to their views on teacher-pupil rapport that they exclude the possibility of success for that teacher.

Now, the differentiation of the concerns of the researcher who intends to help administrative officials select

teachers from those of the researcher who intends to help preparatory institutions select potential teachers indicates that the researcher with the first concern must deal essentially with the question of predicting whether a particular teacher will succeed in a particular job, which, the reader will recall, is the second of the major questions which are the concern of this paper. On the other hand, the researcher who intends to help preparatory institutions is focally concerned with the third and fourth of the major questions, identifying and predicting performance on the problems which characterize elementary school teaching, but do not equally well represent all jobs within that field.

QUESTION 2: How Can One Predict How Well a Particular Elementary School Teacher Will Perform a Particular Job in Elementary School Teaching?

Centrally, the question as to how one can predict how well a particular elementary school teacher will perform a particular job in elementary school teaching involves the problem of matching teacher characteristics to job requirements; that is, the problem of finding those teacher behaviors which are closely associated with success in discharging teacher functions as defined by a particular school system. Solving this problem is a very complex matter on which only a little headway has been made in spite of the immense number of studies in the general field of teacher effectiveness (23: 657; 24: 1490, 9; 13). One reason that the problem is complex lies in the number of variables to be dealt with. As mentioned in the previous section, the definitions of teacher function vary according to grade level and school system. Besides, classroom conditions also vary, with different classes characterized by differences in intelligence, socioeconomic background, previous educational experiences, and the like. In addition, the researcher is confronted with the fact that the procedures that might be used to assess teacher effectiveness do not correlate very consistently with one another (21: 18, 24, 31).

All of these variables make research in teacher effectiveness very complex (22: 261); yet, even though the number of variables with which the researcher in teacher effectiveness must deal is an immense problem, equally significant is the problem of maintaining relevance between the procedures selected by which to carry out research and the objectives of that research. Focally, the problem is one of avoiding the earlier mentioned paradox.

Of the procedures which a researcher might use to assess teacher effectiveness, one, the over-all administrative rating scale, was shown to be of little utility. The procedures which remain differ from the over-all scale in that they contain more structure and, in the case of achievement tests, definite content. The fact that these various procedures, detailed administrative rating scales, detailed student rating scales, systematic observation, and standardized tests, are

structured and in some cases have content is important, because the structure and content of a device for assessing teacher effectiveness represent surreptitious ways of slipping into the criterion measure the researcher's viewpoint about what a teacher's functions should be.

As soon as one states what the rater is to look for in the activity of a teacher, or as soon as one writes items for an achievement test, he has operationally specified precisely what he believes to be the important teacher functions. But it is immediately apparent that, when a researcher loads his own specifications of teacher function into his assessing devices, he has set about the business of re-defining for a particular school system what the functions of its teachers should be. For instance, if the researcher is going to use pupil gain in arithmetic as the criterion, the standardized arithmetic test he chooses to use to assess gains will ultimately define, with great specificity, what functions are to be discharged by the teacher whose pupils he measures. Standardized tests in arithmetic, however, vary widely among themselves, some stressing basic mathematical understandings, others stressing manipulatory skills. If the researcher chooses a test which stresses manipulation to assess teachers of grade 4 in a school system which stresses that teachers convey basic mathematical understandings, he has chosen to define teacher function in that system in a way that is of peripheral relevance to the functions which the system itself holds that teachers should discharge. He has, in other words, re-defined teacher function.

Rating scales or structured systematic observations (those that tell the rater what to look for as opposed to observations which simply record what the teacher does in fact do) define function less specifically than do achievement tests. Yet, a rating scale is likely to reflect with greater or lesser emphasis one of the four points of view about teacher function which, according to W. S. Monroe (20), have developed in the United States beginning about 1890. The earliest viewpoint envisaged teacher function as developing the powers of the mind. In a somewhat later viewpoint (circa 1915) the function of the teacher was centrally to transmit subject matter. More recently teacher function has been conceptualized as guiding learning activities or, in a slightly different viewpoint, directing learning activities. A rating scale constructed by persons whose viewpoint on teacher function centers in the transmission of subject matter is likely to focus on how the teacher organizes and presents subject matter. On the other hand a scale constructed by persons conceptualizing teaching as guiding learning activities is more likely to stress the quality of pupil-teacher interaction and rapport, and to emphasize the skill of the teacher in setting the stage for learning, i.e., in manipulating the learning environment.

To whatever extent either an achievement test or a rating scale re-defines the job of a teacher, the researcher runs the risk of not being able to predict how well that teacher will perform his particular job, since the test or scale used by the researcher as his criterion may not represent the par-

ticular job sufficiently well. It can thus be seen that, in order to avoid the paradox, the researcher must avoid redefining the functions of a particular teacher and at the same time be able to represent the particular set of functions that teacher discharges.

Changing research strategy. It should at this point be clear that in order to predict how well a particular teacher performs his particular job, a change in traditional strategies and procedures in assessing teachers is required. Among the most perspicacious advocates of a systematic strategy has been D. G. Ryans, who has set forth the rubrics of the strategy in a series of papers over the past few years (24; 25; 26; 27). In broad outline this strategy consists of (1) searching for relationships between the classroom behaviors of the teacher and the concurrent and subsequent behaviors of pupils, (2) identifying the specific conditions under which a specified relationship between a pattern of teacher behavior and a pattern of pupil behavior occurs, and (3) demonstrating that a teacher with known characteristics will, in fact, engage in a particular pattern of behavior under a particular set of classroom conditions.

The strategy which the authors have identified with Ryans avoids the paradox but also generates certain problems of its own. By focusing on the relationships between teacher behavior and pupil behavior this strategy avoids having to specify what the functions of a teacher should be. Rather, if a school system can specify what behaviors it believes to be highly desirable in the pupils who attend school in that system, then it becomes possible for the researcher to provide information bearing on the particular type of teachers who bring about such pupil behavior under the conditions which exist in that system. Clearly, this strategy runs comparatively little risk of providing administrative officials with irrelevant, inconsequential, or negatively related information, since there is no attempt on the part of the researcher to redefine the functions of the teachers in a particular school system.

The problems generated by the strategy under discussion are, with one important exception, problems which can be resolved by imaginative endeavor by educational researchers. One of these problems is to create ways in which teacher behavior can be usefully described and classified. An example of this type of research has been recently described by B. O. Smith (30), who has been concerned with describing logical operations in teaching by means of classifying the verbal activity of secondary teachers. This study uses the time sample technique, which consists of either writing down exactly what the teacher says and does or writing down what he does and tape recording what he says. The time sample method has been used in the study of child behavior for many years, but its advocacy for studying teacher behavior has been much more recent (27: 1487).

A second problem in carrying out the strategy under consideration is to create new measures of pupil behavior. In a recent study, for instance, Medley and Mitzel (18: 240) em-

ployed The Russell Sage Social Relations Test, a measure of group problem-solving behavior, and the My Class Inventory, a measure of pupil-teacher rapport from the viewpoint of the students, as measures of pupil behavior. Many aspects of pupil behavior, such as perseverance at various types of school work, avoidance of various types of school work, use of effective inquiry techniques in problem solving, evidences of hostility, degree of mastery of the various developmental tasks of middle childhood, etc., must eventually be objectively measured if "pupil behavior" is to mean more than performance on standardized tests.

A third problem is to demonstrate that particular pupil behaviors are associated with particular teacher behaviors. The difficulties inherent in consistently demonstrating such associations are apparent. Pupils are influenced by many factors other than their teacher. The ways in which the influence of a particular teacher interacts with these other influences in the child's life are complex and difficult to tease out, even in well designed studies using refined statistical techniques. Whether clear-cut, significant relationships between teacher behavior and pupil behavior can ever be found is a focal question.

A fourth problem which is inherent in the strategy which the authors attribute to Ryans is to show that teacher behaviors assessed before the teacher takes a particular job can be used as predictors of his behavior while carrying out that job. At the present time, this problem is a hypothetical one, since there is not much point in trying to predict that a particular pattern of teacher behavior will occur unless it is known that that particular pattern is significantly and consistently related to a particular kind of pupil behavior under a particular set of conditions. How pupils respond to a particular pattern of teacher behavior as conditions vary is, for all practical purposes, unknown.

Appraisal of the change in strategy. It may be said at the outset that the ultimate test of a strategy directed toward predicting how well a teacher will perform his particular job is whether it leads to such predictions. Since the newer strategy has been recently developed, any discussion of its efficacy in this respect is premature. On the other hand, every strategy involves assumptions, and these may be critically examined as soon as they appear. The intent of the authors at this point is to examine one of the assumptions which appears to underlie much thinking about changes in strategy.

One of the most difficult problems in the study of teacher effectiveness has been whether to assume that "effectiveness" is a statement about an attribute of the teacher, whether it is a statement about an attribute of a teacher in a particular teaching situation, or whether it is a statement about the results which come out of a teaching situation. The first Report of the Committee on the Criteria of Teacher Effectiveness (22) stands as a singular monument to indeterminateness on this assumption. At the outset the committee appears to think that effectiveness is a property of the teacher (22: 242-243) but, by the Second Report, the committee seems

more prone to think that effectiveness is a "now you have it, now you don't" property of the teacher, depending on the situation (23: 644-657). Ryans has taken the position that effectiveness is partly dependent on the attributes of the teacher and partly on the teaching situation.

The difficulty with what assumption to make arises, of course, because neither the assumption that effectiveness is a property of the teacher nor the assumption that effectiveness is a statement about a teacher in a particular situation has been tested. There seems to be the feeling that most studies in teacher effectiveness are searching for a property of the teacher (1: 266) and that this search has not been successful (23: 657). Yet, to the knowledge of the authors, the assumption on which this search is based has not been tested, for, in order to test it, it is necessary to measure the results of a teacher's behavior on constant criteria as that teacher is confronted with many different classroom situations in many different school systems. To test adequately the assumption that effectiveness is a property of the teacher requires a longitudinal study with the repeated measurement of the same teacher on the same criteria under a wide variety of teaching conditions. Such a study, to the knowledge of the authors, has not as yet been conducted.

In spite of the fact that effectiveness as an attribute of the teacher has not been shown to be untenable, there appears to have been a general espousal of the term "effectiveness" as a statement about the results of the behavior of a teacher in a particular teaching situation; that is to say, whether or not a teacher is effective depends on the variables operating in the situation he is in.

In examining this assumption it is best to begin by recognizing that the assumption actually lies along a continuum of such assumptions. At one end of the continuum is the assumption that effectiveness is determined almost wholly by the teacher, can be viewed as one of his attributes, and depends very little on the variables operating in the situation; while at the other end it is the assumption that effectiveness is almost wholly determined by the particular variables operating in the situation in which teaching occurs. While the researcher is free to choose his assumption from any place on the continuum, in choosing it he implicitly makes a hypothesis about the adaptability of teachers to teaching situations. If a researcher proceeds on the assumption that effectiveness is almost wholly dependent on the particular variables operating in a situation in which a teacher teaches, he seems to be implicitly hypothesizing that teachers simply emit behavior without close regard to the appropriateness of that behavior with respect to the variables operating in the situation. If a teacher happens to emit the right behavior for a particular set of conditions, then he is effective, or, in terms of Ryans' strategy, his pattern of behavior leads to certain (desirable) pupil behaviors under those conditions.

At the other extreme, if a researcher proceeds on the assumption that effectiveness is almost wholly attributable

to the teacher, he implicitly hypothesizes that teachers are adaptive in their teaching behavior so that they emit behavior with a close regard to its appropriateness, or, to put it in another way, that teachers are capable of fine discriminations in their environment and have available the responses appropriate to those discriminations. In this view, teachers are *effective* because they adjust their responses to the variables operating in a situation in order to bring about desired pupil behavior.

It is apparent at this point that a choice of one assumption or the other is not desirable, because the act of choosing, itself, makes it necessary to exclude one of the assumptions, and there is no reason to think that effectiveness necessarily depends on variables operating in the situation or that it necessarily depends on the teacher. It may be hypothesized that both assumptions are valid ones, depending on differences among the sub-populations of teachers. Some teachers are no doubt capable of producing desirable pupil behaviors in a wide variety of teaching situations. The behavior they display, of course, depends on the situation, but this is not the point; the point is that their behavior is appropriate to that situation and that desirable pupil behaviors result because it is appropriate. These teachers do not have some sort of magical property called "effectiveness," but they are characterized by adaptability to teaching situations or, as the authors prefer to put it, they have developed a high level of skill in dealing with the problems which arise in the course of their professional work.

There are some teachers, on the other hand, who are no doubt capable of producing desirable pupil behaviors under certain teaching conditions but not under others. These teachers are *effective* within the limited range of teaching situations for which the responses they have available are appropriate. This group includes all those teachers whose values, attitudes, scope of knowledge, and teaching methods fit well into the values, attitudes, knowledges, and views on teaching methods expressed by the persons living in the community in which they teach. These teachers can presumably be successful in all communities highly similar to the ones in which they are teaching successfully.

Finally, there are no doubt some teachers whose patterns of teaching behavior are so poorly developed that they can succeed in only a very few teaching situations or not at all.

If these hypotheses bearing on various sub-groups of teachers are valid, there is an implication which bears on the uses of the results of the newer strategy.

While the results of the application of the newer strategy should be highly useful for matching teachers to particular jobs, there is an inherent danger of mis-selecting and mis-preparing preparatory teachers unless the results are used with great care. If, for example, one finds that a teacher with a particular pattern of teacher characteristics is associated with *effectiveness* in a particular type of

teaching situation, there is an implication that some preparatory teachers with this pattern of characteristics should be selected to function essentially as effective replacements. This seems reasonable, but, if preparatory teachers are selected and prepared primarily to function in a particular type of teaching situation, a question arises as to what will become of the effectiveness of these teachers as teaching situations change, as they surely will unless the social modifications within the nation suddenly cease.

QUESTION 3: How Well Does a Particular Elementary School Teacher Solve Those Problems Which Characterize Elementary School Teaching as a Professional Endeavor But Which Do Not Equally Well Represent All Jobs in Elementary School Teaching?

In order to avoid both the paradox mentioned in connection with Question 1 and the difficulties which the authors believe to be inherent in the newer strategy, it is ultimately necessary to change the objectives of research on how well teachers perform and to evolve a conceptualization of teacher behavior which will be distinct and clear in its assumptions and will lead to fruitful results and ultimately to the increased systematization of knowledge about teachers and teaching.

As long as the objectives of research concerned with teacher performance are directed toward attempting to predict how well a teacher is going to perform in a particular situation, the researcher must make the variability in such situations one of his major concerns. He may, in fact, make it a concern of such magnitude that it obscures other, perhaps more significant, considerations. One of these considerations is the socially controllable conditions under which teaching is expected to occur. In the school system mentioned in an earlier illustration, teaching was to take place under very rigid conditions, where there was, in effect, no professional autonomy for the teacher. The behavior of such a teacher might be extremely restricted, either because he has only a narrow range of responses available while teaching or because he sees the restriction of his behavior as a condition of holding his job. A researcher might be interested in predicting what characteristics are necessary to succeed in such a job, or he might ask why certain teaching situations require a particular pattern of teacher characteristics for success.

It is the case, then, that two perspectives on the relationship between teacher behavior and teaching conditions are possible. In one perspective the researcher focuses on the situation in which a teacher teaches without consideration of the socially controllable conditions which limit the teacher's behavior in that situation. He then asks what teacher characteristics are associated with success on the criterion of desirable changes in pupil behavior in that situation. The advocates of the newer strategy seem to take

this perspective. In the other perspective, the researcher focuses on teacher characteristics and asks what characteristics or what degree of a particular characteristic (intelligence, for instance) tends to be associated with teachers who teach in a particular school system, with full recognition that a particular school system may define the conditions in such a way that teachers with particular characteristics, such as high intelligence and a wide scope of knowledge, are driven away from the system.

The authors therefore take the position that the number and kind of changes a teacher makes in pupil behavior are not simply the results of teacher behavior, but are the results of teacher behavior when some kind of limits are placed on the behavior permitted to the teacher. It is quite thinkable that the teachers who could produce the best results in a particular school system either cannot be attracted there or are driven out by administrative policies, community practices, and social pressure from teachers who do not produce the best results but who are well adapted to administrative policy and community practices.

The second of the two perspectives implicitly involves the hypothesis that particular segments of the teacher population tend to gravitate toward particular kinds of teaching situations. This hypothesis is based on two assumptions: first, that school systems do not compete equally in the teacher supply market. Some school systems have a great deal more to offer teachers in the way of teacher autonomy, community stature and progressiveness, location, and salary, than do other systems. Second, it is assumed that school systems tend to select and, more significantly, to retain teachers who are in general accord with the prevailing views on educational practice in that system.

If one should pursue this hypothesis and these assumptions far enough in one direction, he might eventually engage in matching teacher characteristics with particular teaching jobs. But there is another direction that can equally well be taken, not to the exclusion of the first direction, but prior to it. This is to examine the kinds of characteristics that appear to be focal to teaching as a learned skill and to construct instruments to measure such characteristics.

At this point some legitimate questions about strategy can be raised. To construct instruments which are focally related to teaching skill assumes that one knows something about it. It has been one of Ryans' points that much thinking about teaching and, particularly, effective teaching is essentially a priori, i.e., it rests on unestablished assumptions. The authors have no quarrel with the observation that much thinking about effective teaching is a priori, but they do diverge from Ryans' position with respect to the strategy for remedying this state of affairs. The strategy which the authors attribute to Ryans is a maximum security strategy. It involves no risk; it inherently involves no theory although theory may be brought in from external sources (26); and it accumulates information slowly and in an unsystematic way.

The older strategies in the exploration of teacher effectiveness were also maximum security strategies, involving no risk, clearly involving no systematic theory, and unfortunately yielding virtually no useful information. The "useful" information yielded by over 1,300 studies on teacher effectiveness is summarized in the third edition of the Encyclopedia of Education Research (19: 27) in approximately three quarters of one column, although a total of approximately 17 columns is devoted to discussion of teacher effectiveness.

It seems essential at this time to adopt a strategy which involves a considerable amount of risk and an inherent theory (a part of the risk taken is, of course, that the theory will be wrong) but which will yield systematic information about teacher performance. From a short, practical viewpoint, the central objective of the strategy is to provide information upon which decisions involving the selection of preparatory teachers can be made. From a somewhat longer viewpoint, the objective is to collect and systematize information about teacher performance on controlled problems in teaching and about the behavioral attributes which are associated with this performance.

Toward the development of a theory as an implicit part of a strategy in the study of teacher performance. It is apparent that the development of a strategy involves making postulates or hypotheses and then selecting procedures consistent with the postulates and hypotheses that are consistent in the sense that they do not change the objectives of the research or fail to test the hypotheses. One of the assumptions which has already been stated and which is central to the present strategy is that the population of teachers consists of sub-populations of teachers who have different characteristics. Of the sub-populations, one group is hypothesized to have characteristics which permit its members to succeed in a wide variety of teaching situations, while the other groups have characteristics which permit them to succeed in relatively few situations.

In order to have some idea of what these characteristics are and why they enable that particular group of teachers to succeed, hypothetically speaking, it is necessary to elaborate several interrelated postulates and hypotheses about what kind of behavior teacher behavior is. First it may be postulated that teacher behavior consists of two kinds of responses: (1) conditioned responses and (2) instrumental responses. In the theory which the authors have in mind, conditioned responses are important only to the degree that they interfere with or facilitate instrumental responses.* The focal behavior

---

*The authors hope to develop the role of conditioned responses in teaching at a later time. Because so little is known about the instrumental responding of teachers, it is difficult to state any but the most general hypotheses about conditioned responses in relation to these. From the work done in academic psychology, however, it is apparent that such factors as conditioned avoidance can play a considerable role in the behavior of a teacher.

of the teacher in the classroom, then, consists of responses which are instrumental, or intended to be instrumental, in bringing about a goal state which, in teaching, consists of a pupil response or a set of pupil responses which are held to be desirable by a particular teacher.

It is possible to talk about the teacher's behavior in somewhat different terms. When a teacher sets for himself the goal of bringing about a particular pupil response or set of responses, he generates a problem for himself. The problem is how to get the pupil to make the response the teacher wants him to make, and then to make certain that the pupil can make the response when he needs it to help him solve the problems that will confront him both as a pupil and as an adult. Teacher behavior, then, is problem-solving behavior of a particular kind.

Ryans (26), Sears (28), and Tiedeman and Cogan (32) have all presented theoretical papers which view teaching behavior as involving instrumental responding, but they have not viewed teacher behavior as problem-solving behavior or followed the model through to total development or empirical testing. A second rather curious observation is that studies in problem-solving skill as a predictor of teacher effectiveness are not, to the knowledge of the authors, to be found in the literature--a fantastic state of affairs for a group which talks as much about problem solving and thinking as educationists do.

Following up the notion that teacher behavior is problem-solving behavior, it is apparent that a group of teachers who could succeed in a wide variety of teaching situations would consist of persons who were good solvers of teaching problems--wherever such problems appear. Moreover, such a group would possess certain characteristics, one of which is that its members could solve a wider scope of teaching problems than could the teachers in other segments of the teaching population, although they might not solve a given problem any better than some teachers from other groups. To determine why these teachers might solve a wider scope of problems than other segments of the teaching population, it is necessary to return to the assumption that teacher behavior consists largely, but by no means entirely, of instrumental responding.

Instrumental responses may be viewed as learned responses. If teacher behavior consists predominately of instrumental responses, then it is learned behavior. This notion, together with the notion that teacher behavior is problem-solving behavior, suggests the hypothesis that the teachers who are the best solvers of teaching problems are those who have acquired the greatest number of instrumental responses and who can best transfer those responses to "new" (or variant) situations. If this hypothesis is tenable, then the characteristics of this group of teachers are positively related to the acquisition of instrumental responses and the transfer of these responses.

One characteristic which would be hypothesized to be positively related to the acquisition of instrumental responses would be learning ability. This term is not precisely coter-

minous with measured intelligence. As Gulliksen (15: 515) has pointed out, measured intelligence is only moderately related to learning ability as measured by laboratory experiments. On the other hand the concept of intelligence is relevant to the concept of learning ability, and one could reasonably expect the best solvers of teaching problems to be characterized by greater intelligence than other segments of the teaching population, other things being equal. The trouble is that other things rarely are equal. Intelligence, or, more exactly, learning ability, may be a condition of being a good solver of teaching problems, but its significance may be overshadowed by or confounded with other variables, making its role difficult to determine.

A second characteristic which may be hypothesized to be positively related to the acquisition of instrumental responses relates to the number and type of opportunities a teacher has had to make such responses. The considerations in this respect are complex and need to be carefully sorted out. First, in order to acquire a broad repertoire of instrumental responses which will transfer to many kinds of teaching problems, it is essential that a teacher at some point be permitted to make such responses. However, in situations where teacher autonomy is severely limited, particularly in the first years of teaching, the opportunity to acquire a wide range of instrumental responses is also severely limited. In school systems where a single method of instruction is advocated, where a very limited set of objectives is sought, where individual differences are suppressed, and where these practices are rigidly enforced by social pressure from other teachers and by direct pressure from administrators, the teacher has little opportunity to learn a wide range of instrumental responses.

It is therefore hypothesized that the best solvers of teaching problems tend to be found teaching in school systems where they enjoy a broad professional autonomy. This hypothesis is made not simply because autonomy for teachers is a "good idea," but because it is a condition of learning that in order to acquire a response one must first have the opportunity to make it.

A second matter with respect to the opportunity to acquire a wide range of instrumental responses bears on the experiences of the teacher before entering teaching, such as working with boy scouts or girl scouts, but the problem with such situations is that it is not possible to tell whether the responses learned will transfer positively or negatively to problems in teaching. Both teaching and being a boy or girl scout leader involve working with children, but how many other similarities there are between the two depends greatly on the particular objectives of the organizations.

A second type of pre-service experiences which provide opportunities to learn instrumental responses which will transfer to the solution of teaching problems are those in the professional education curricula of teachers colleges and schools of education. It is immediately apparent that the opportunity a pre-service teacher has to acquire instrumental

responses in professional education depends on the college he attends. It is therefore necessary to determine what kind of professional education is positively related to being a good solver of teaching problems rather than to assert that such education is positively related. Nonetheless, because it has already been hypothesized that the amount of opportunity to acquire instrumental responses is positively associated with being a good solver of teaching problems, it must also be hypothesized that those professional education curricula which provide the greatest opportunity for pre-service teachers to make instrumental responses with children toward a wide range of goals contribute proportionally more persons to the group of good problem solvers than do other curricula, other things being equal.

A third characteristic which may be hypothesized to be positively related to being a good solver of teaching problems is related to motivation. Again the matter is a complex one. As a general concept, motivation may be split into two dimensions. Along one dimension one may talk about the type of motivation, e.g., hunger, thirst, anxiety, achievement, "effectance" (36), and so on. How profitable it is to talk about particular types of motivation, particularly by means of such concepts as "drives," is a focal question in contemporary psychology. The serious reader may want to refer to Bindra's recent book (4) for an excellent discussion of this question. The authors prefer to follow Bindra's viewpoint and to discuss type of motivation by reference to observable incentives rather than to drives.

Along a second dimension one may talk about level of motivation, i.e., how strongly a person is impelled toward a particular goal object or incentive. In order for a teacher to acquire a wide range of instrumental responses which transfer to a wide variety of teaching problems, it may be hypothesized that at least a moderate level of motivation is requisite. It may further be hypothesized that in order for a wide range of instrumental responses to be acquired, this level of motivation must have considerable continuity. If these hypotheses are tenable, the next problem is to attempt to identify what types of incentives are relevant to maintaining moderate levels of motivation. The authors hypothesize that, while many types of incentives enter into being a good problem solver, one type must be positively related to being a good solver of teaching problems.

Instrumental responses by definition are those which bring about some goal or goal object for the responder. In teaching, the goal is to change pupil behavior, to be able to see a pupil make a new and desirable response as a result of teaching. If the consequences of instrumental responses are a condition of their acquisition, as is indicated by Skinner (29), Bindra (4), and Sears (28), then perceiving changes in pupil behavior must be a central incentive in the acquisition of the instrumental responses which compose teacher behavior. Moreover, perceiving changes in pupil behavior seems to be the only incentive in teaching which comes directly at the end of instrumental responding and which could definitely function as

a reinforcer. (By a "reinforcer" is meant a consequence of the instrumental responses of the teacher which increases the probability that those instrumental responses will be made again under the same or similar circumstances.)

If it is true that the perception of desirable changes in pupil behavior is a central incentive in the acquisition of the instrumental responses which make up teacher behavior, then the group of teachers who have acquired the largest repertoire of instrumental responses should have two characteristics. First, they would be sensitive to a wider range of educational objectives (desirable pupil behaviors which can be brought about by teaching) than are other segments of the teaching population. One of the reasons that good solvers of teaching problems can solve a wider range of problems in teaching is that they see a wider range of educational objectives as desirable and learn how to respond toward the attainment of these objectives. Second, the perception of changes in pupil behavior would have a greater incentive value for good solvers of teaching problems than for other segments of the teacher population. Teachers who solve teaching problems well enjoy seeing youngsters grow in their attainments.

The reason that a particular group of teachers would respond toward a wider range of educational goals and find greater incentive in these goals than other groups no doubt lies partly in the early experiences of the members of the group of teachers who are good solvers of teaching problems; however, it may be hypothesized that it also lies in the range and type of educational goals emphasized as important and learned as desirable and valuable in the teacher's liberal arts and professional education. It has been hypothesized earlier that those professional curricula which provide the greatest opportunity for preparatory teachers to make instrumental responses with children toward a wide variety of goals contribute a greater proportion of teachers to the group of good problem solvers than do other curricula. To this hypothesis may be added a second, somewhat more general, hypothesis, that membership in the group of good solvers of teaching problems is a function of the scope of objectives emphasized as important and learned as desirable and valuable by students in the preparatory institutions from which teachers come, such that the wider the scope of objectives emphasized by the preparatory institution and learned by students, the greater the proportion of teachers it will contribute to the group of good solvers of teaching problems, other things being equal.* These two hypotheses generate a third; namely, that preparatory institutions which emphasize a wide range of objectives which are in turn learned by students and which provide ample opportunity for preparatory teachers to make responses with children contribute more teachers to the group of good problem solvers than do other institutions.

---

*That the general or liberal arts background of the student makes substantial contributions to the goals a preparatory teacher learns as valuable goals in education is implicit in this hypothesis.

The hypotheses made so far concern some of the characteristics of teachers who have been hypothesized to acquire a wider range of instrumental responses relevant to a wide range of educational objectives, and teacher behavior as problem-solving behavior has been left somewhat on the periphery. If individual pupils and classes of pupils were always exactly the same, teachers who acquired a very great number of instrumental responses to a wide number of goals would have no problems. They would have learned exactly what responses to make to bring about a particular change in pupil behavior. However, individual pupils vary, and so do classes. As every experienced elementary teacher recognizes, how "easy" or "difficult" a class is depends in a large part on what particular combination of individual pupils appears in that class. One retarded pupil, one gifted pupil, and one emotionally disturbed pupil can, in combination, completely upset what would otherwise have been an "easy" class. Individual pupils vary in many ways, not only in achievement and in the rate at which they can acquire new responses, but also in the ease with which they can be stimulated to respond toward educational objectives, in the ease with which they adapt to other class members, and in emotional stability. Teachers acquire problems, then, not only because they hold certain objectives desirable, but also because the conditions under which these objectives are to be attained vary greatly.

In discussing the acquisition of instrumental behavior as relevant to problem-solving behavior, it has been implicitly hypothesized that the greater the number of instrumental responses a teacher has available, the greater the probability that he can solve a particular problem. This hypothesis is made on the grounds that the more responses an individual has available, the more likely it is that at least a few of these responses will transfer in some degree to a "new" or variant situation--a situation which is not exactly like one he has encountered before.

There are, however, some other conditions besides transfer attached to solving a problem. This is to say, even though a person has a great many responses available, he still may not be a good problem solver. This point is worth attending because it means that, while all good solvers of teaching problems are hypothesized to have many responses available, not all persons who have many instrumental responses available are hypothesized to be good problem solvers. Having responses available is a necessary but not a sufficient condition of solving teaching problems. What, then, are the other conditions?

A number of different perspectives are possible on what is required in solving a problem besides having several instrumental responses available. Following a perspective suggested by Bindra (4: 57), what is required is for the person to discriminate the variations in the stimuli which surround the goal and then to select responses accordingly, i.e., to respond "appropriately." In the behavior of a teacher this would mean that what is required is for the teacher to identify exactly what it is about the particular situation in

which he is attempting to attain a given goal that makes it different from other situations in which he has attempted to attain this goal, and then to make a response appropriate under the changed conditions. The problem is not to make totally "new" responses, but to select from among old responses or perhaps to modify old responses slightly or combine them in a new way.

In a somewhat different perspective than the one just given, the matter of the teacher selecting an appropriate response may be seen in a context in which the teacher risks that one of the responses or one of the combination of possible responses will obtain better results than the others, that is, will result in greater pay-off (6: 185). The pay-off in this case is the degree to which a given pupil or group of pupils, according to the particular circumstances, attains the behaviors which the teacher holds desirable for that pupil or group. The greater the degree of attainment, the larger the pay-off. The problem for the teacher is to make as certain as he can that the set of responses he makes will be the set that brings about the greatest pay-off for the time and energy he has available to expend on the attainment of the objective he has in mind. If this particular perspective on problem solving is followed through, certain aspects of what teachers might do in solving teaching problems come into sharper focus.

In order to make as certain as possible that one set of responses rather than another will lead to the greatest pay-off, the teacher problem solver must undertake a search for information which will increase the probability that a particular set of the responses will produce the greatest pay-off. In conducting this search the teacher problem solver has to be able to do at least two things: (a) he must be able to generate some sort of guide or "search model" which serves to direct investigation toward relevant information, and (b) he must judge or "weigh" the significance of the information (i.e., accurately utilize the information) once he obtains it. Exactly how a person conducts his search may be called his "search pattern" or his "strategy."

What is significant at this point is that it is known from other studies, such as that of Bruner, Goodnow, and Austin (6) and that of Bryan (7), that there are wide differences among individuals in the strategies they adopt for solving problems of a particular kind. Some strategies are, of course, more optimal than others. A proficient problem solver is presumably one who adopts an optimal strategy each time he confronts a problem. What is hypothesized, therefore, is that, in addition to having many instrumental responses available, a good solver of teaching problems is a person who is capable of adopting an optimal strategy for ascertaining which of the possible instrumental responses will lead to the greatest pay-off under a particular set of conditions.

What are the characteristics, then, that may be attributed to teachers who adopt optimal strategies to ascertain which instrumental responses will result in the greatest pay-off? The authors do not know the answer or answers to this question, nor do they know what kinds of strategies teachers

characteristically employ in solving their problems. They do not even know whether teachers as a whole do engage in problem-solving activity, but they do hypothesize that the good solvers of teaching problems do engage in such activity.

The development of procedures and criteria as an implicit part of a strategy in the study of teacher performance. In order to identify the good solvers of teaching problems, it is clearly necessary to find or develop teaching problems for teachers to perform. Such problems should exist in the classroom, and it therefore seems desirable to study the problem-solving behavior of teachers in that context. However, such a procedure seems to be questionable in view of the objectives for which it would be done. There are three reasons why the authors believe it questionable.

First, if good solvers of teaching problems are those teachers who can solve a wide scope of teaching problems under a wide variety of conditions, then to identify such a group would require (a) moving each of a group of teachers being studied to a different set of conditions every year or every other year, and (b) studying this group of teachers on a longitudinal basis, i.e., following them year after year to situation after situation. Such manipulation and control of teachers is neither possible nor desirable from any other than a research point of view. Moreover, the very fact that such a group would have to consent to be studied and would recognize that they were "participating" might significantly influence their behavior.

A second reason for not studying the problem-solving behavior of teachers in the classroom is an extension of the first. In order to make comparisons of skill in solving teaching problems, it is absolutely crucial that all teachers be examined on the identical problems and over the same scope of problems. There is not much point in attempting to compare how well a teacher succeeds in solving problems in the third grade in certain areas on Chicago's South Side with how well a teacher solves the problems in a third grade in Winnetka. The problems are not comparable. To fail to recognize their non-comparability would be to commit the same fallacy as contained in the conclusion that "two teachers who are 90 per cent effective are effective in doing the same things." To accept this proposition is to commit a formal logical fallacy, viz., the fallacy of accident (10: 377).

A third reason for not studying the problem-solving behavior of teachers in the classroom lies in the difficulties attendant upon demonstrating that a change in pupil behavior has occurred because of the instrumental responding of a single teacher rather than because of other factors. The central difficulty is one of establishing sufficient experimental controls to show that certain changes in pupil behavior occur if and only if these changes are preceded by the instrumental responding of a given teacher. To attain precise controls in the classroom is probably impossible, although various statistical controls, including matching groups, analysis of covariance, and randomized designs, serve as approximations. A

difficulty linked to the matter of demonstrating that particular pupil behaviors are attributable to the instrumental responding of the teachers lies in the amount of confidence that can be placed in the criterion measure of pupil behavior. For instance, the immediate goal of a teacher's instrumental responding may be to get his pupils to perform long division problems of a given level of difficulty. If his pupils learn to perform such problems, one has some confidence that he has solved the problem; but if at some later time these same pupils have difficulty learning to do more difficult problems in division, or have difficulty learning to use division in algebra, or cannot use division to solve elementary science problems, confidence in the initial criterion is greatly shaken.

At this point a disturbing state of affairs appears. As one moves from the more immediate and more convenient criteria for assessing how well a teacher solves his problems to those which are more distant, but which are more valid (from the viewpoint of why a teacher tries to bring about new and desirable pupil responses in the first place), the pupil behavior attributable to a given teacher becomes increasingly intertwined with the behavior produced by other teachers. Stated in another way, as one's confidence in the criteria of effectiveness increases, the probability of being able to attribute this behavior to any one teacher correspondingly decreases. The reason why this is the case is not far to seek. The more distant criteria for assessing the effectiveness of a teacher or his problem-solving proficiency are criteria which depend on the transfer of training. Transfer phenomena (proaction) depend both on initial learning, or what is usually called the training trials, and on what is to be learned or done later, i.e., the "test" trials. Thus, one of the criteria for how well a fourth grade teacher teaches arithmetic depends on how well his students learn fifth grade arithmetic, while one of the criteria for how well a fifth grade teacher teaches arithmetic is the extent to which he can create situations to which the students coming out of the fourth grade can transfer their knowledge. The fourth and fifth grade teachers therefore become interacting units, and to which of them the arithmetic behavior of the students at the end of grade 5 should be attributed becomes a very interesting problem in logic and in statistical analysis--a problem which, along with those previously mentioned, suggests to the authors that studying the problem-solving behavior of the teacher might better be done under non-classroom conditions.

In order to study the problem-solving behavior of teachers outside of the classroom it is necessary to do three things: (a) develop problems, (b) develop a way to assess performance on the problems, and (c) show that the problems are useful for talking about the problem-solving behavior of teachers.

In the development of teaching problems three conditions should be met. First, the problems should be such that all the variables which enter into them can be carefully controlled. Each problem should be exactly the same for each person who attempts to solve it. Second, the problems should represent, as nearly as possible, problem situations which ele-

mentary school teachers might encounter in teaching. Third, the problems should be such that some of them are directed toward assessing the scope of problems a teacher can solve, while others are directed toward the investigation of the strategies which teachers employ in solving teaching problems. Problems directed toward assessing the scope of problems a teacher can solve should be comparatively short, involving moderate amounts of performance time, easily administrable and easily scorable. Problems directed toward the investigation of strategies, on the other hand, are essentially laboratory problems, involve long performance times, and require complex scoring schemes.

For a full development of the procedures which enter into problem development, the reader should consult two other publications by the authors (33; 34). For the purposes of the present paper, however, one aspect of problem development is focal--the selection of the educational goals which serve as guides in problem construction.

As stated earlier, a problem arises for a teacher when he assumes a particular educational objective with a single pupil or group of pupils. The problems a teacher must solve, then, depend upon the educational objectives he assumes as well as upon the conditions under which he assumes them. Moreover, the particular function a teacher discharges or attempts to discharge is explicitly defined by the objectives he assumes. What a particular teacher's job actually consists of, then, is determined by the educational objectives which he freely assumes or is compelled to assume by authorities in the school system.

In the construction of teaching problems, however, only a limited and logically consistent set of educational objectives can be assumed. For instance, the set of arithmetic objectives assumed in the problem construction undertaken by the authors (34) excludes that of having pupils be able to perform division by the short division algorism in grade 4. If a teacher freely assumes or is compelled to assume this objective in teaching grade 4, there is no means by which that particular aspect of his particular job can be represented in the problems constructed for fourth grade teachers, since this objective is excluded at that level.

It first appears that the problem of representing teaching jobs would be slight, but a little further thought suggests that, if the objectives teachers work for were placed on a frequency distribution or a probability curve, those in some jobs would be very deviant or atypical. For instance, data collected by one of the authors indicate that teaching for understanding of the fractional-part meaning of division is rather atypical in the middle grades of the elementary school, even though this meaning of division underlies comprehension of common fractions, ratios, and percentages. Clearly, no matter what set or sets of goals are selected, some jobs will not be represented. The question then becomes one of choosing those goals which are held to be desirable either after a systematic study of desirable goals in the elementary school, such as those published by Kearney under the auspices of the

Russell Sage Foundation (17) or those published by authorities in particular areas of the curriculum; for example, those published by Brownell in arithmetic, under the auspices of the National Council of Teachers of Mathematics (5), by Greene and Gray in reading and language arts, under the auspices of the National Society for the Study of Education (14), and by Havighurst in the developmental tasks of middle childhood (16).

Increasing the number of goals upon which problems may be constructed increases the scope of the problems, but it does not necessarily increase the representativeness of the problems. It probably tends to increase the representativeness of the problems for teaching situations in which the functions of the teacher are regularly re-defined to keep abreast of increases in knowledge and developments within the profession, and to decrease the representativeness of the problems for teaching situations in which function is controlled by local traditions.

The selection of goals has a direct consequence in the identification of variables which will affect teacher performance on constructed teaching problems. As the discrepancy between the goals toward which a teacher characteristically works and the goals on which constructed teaching problems are based increases, the level of performance on constructed teaching problems may be hypothesized to decrease correspondingly. The goals the teacher holds desirable and/or works toward then become a significant source of variation in the problem solving performance of the teacher.

The second major problem in the study of teacher performance on constructed teaching problems is to develop ways of assessing or scoring teacher performance. The difficulties in this area arise because in constructed problems there can be no direct appeal to observe changes in pupil behavior, although indirect appeal to this criterion may be made. If one goes back to the notion that a good solver of teaching problems is one who makes "appropriate" responses, it becomes apparent that there is a way to assess teacher responses without direct reference to pupil behaviors. The "appropriateness" of a response is basically determined by the amount of information available to the teacher up to the point where the response is actually made. If a teacher systematically ignores the information available or is incapable of processing it, the probability of making an appropriate response is correspondly reduced. This does not mean that teachers who ignore or cannot process the information available produce no changes in pupil behavior, it simply means that such teachers cannot control the changes they produce. Their teaching occurs by chance and must be viewed as essentially non-rational.

The assessment of teacher performance on constructed problems is made possible, then, by controlling the amount of information available in the problem. If the information is closely controlled, it becomes possible to specify which responses are most appropriate in that situation. There are, however, differences in problems with respect to how closely the information can be controlled and in the determinateness

of the inferences possible from that information. In some problems, for instance in the diagnosis of errors in performing arithmetic examples, the amount of information available to the teacher can be very carefully controlled and the inferences possible from this information can be held almost completely determinate. In other problems, however, there is much less determinateness of the inferences possible from the information presented, and it is necessary to appeal to research evidence that a particular response or set of responses has the greatest probability of succeeding in that situation. The extent to which an appeal can be made to evidence outside the problem itself is, of course, limited to the availability of such evidence.

The scoring rationale just outlined applies primarily to those teaching problems which are comparatively short and are useful for determining the scope of the problems to which a teacher can respond appropriately. In longer, laboratory problems the focus is partly on whether the teacher's response is appropriate, but primarily it is on how he deals with the information he has available--the strategies he adopts. Since each such laboratory problem actually constitutes a self-contained experiment in problem solving, and since such problems vary widely in the kind of responses they require, the authors believe it most desirable to report the procedures for scoring each problem as the results are published rather than to make an abstract description of these procedures at this point.

The third major problem in the study of teacher performance is to demonstrate that constructed problems in teaching are useful. Usefulness depends on purpose. It is apparent, for instance, that constructed teaching problems, particularly of the laboratory type, could be useful for studying strategies in problem solving and for developing theoretical concepts without having to be useful for other purposes. However, within the purposes developed in the present study, constructed problems must be useful for sorting out, from the general population of teachers, those teachers who can succeed over a wide variety of teaching situations. This is simply to say that the problems must be valid for this purpose.

In the case of constructed teaching problems, the matter of demonstrating validity is a complex one. The best single criterion is to be able to identify, by use of constructed problems, that group of teachers who, when moved from one set of teaching conditions to another, will continue to be able to obtain unusually good results with children. It is not possible to use this criterion, however. The problem therefore becomes one of developing other criteria in which confidence can be placed. It is possible that no single one of these criteria will draw as much confidence as the criterion above, which cannot be used, but it is distinctly possible that the use of several carefully developed related criteria will draw as great or even greater confidence than any single criterion. In developing these criteria, the authors utilized, in general outline, the approaches suggested by Gulliksen (15). The suggestions by Cronbach and Meehl (12) on construct validity, and

the counter-suggestions on construct validity by Bechtoldt (3), are also considered relevant, however.

As Gulliksen has suggested (15: 515), if a test of an aptitude or skill is intrinsically valid, coaching on the test will not only improve test performance but will also improve criterion performance. If constructed teaching problems are valid, then teaching teachers how to solve such problems will, in fact, increase skill in solving classroom problems. Clearly, when coaching on a test invariably improves criterion performance, a great deal of confidence may be placed in the validity of the test. However, in order to use this type of criterion for the validation of a test of problem-solving performance in teaching, it is crucial to have accurate and reliable measures of the problem-solving skill of the teacher in the classroom. Such criteria are at present lacking, but it is believed they could be developed, perhaps, by studying the problem-solving behavior of teachers under laboratory conditions.

While awaiting the development of criterion measures of a teacher's problem-solving behavior in the classroom, an approximation of the criterion suggested by Gulliksen may be employed. In the main, the most rapid changes in the classroom problem-solving skill of teachers would be expected to occur during the first years of teaching, since it is during this time that the frequency with which the teacher is called upon to solve problems increases rapidly. Correspondingly, a valid measure of problem performance would be expected to be sensitive to changes occurring during this time. This is to say that, while it is not possible to show that changes in criterion performance result from coaching on test problems, it is possible to show what is essentially the reverse relationship, that the changes which would be expected to occur in criterion performance in the first years of teaching are reflected in test performance. Whatever it is that causes changes in criterion performance, then, also causes changes in test performance. This criterion does not draw as much confidence as the criterion suggested by Gulliksen, but it does draw some confidence. The degree to which it draws confidence may be best observed by considering how valid a test of problem-solving skill in teaching is which fails to reflect any changes in performance during the period when a teacher would be expected to most rapidly acquire skill in solving teaching problems.

Confidence in this initial criterion can be increased by adding a second criterion. What one wants to demonstrate by using the first criterion is that performance on constructed teaching problems improves because of teaching experience. If this causal relationship is truly demonstrable, then persons who graduate from college but do not enter teaching should not improve their performance on constructed teaching problems, but, rather, should remain constant or decrease in performance. If they do improve, then confidence in the validity of the test is weakened.

A third criterion, which may be seen as an extension of the first and second, is that a test of constructed problems which does not distinguish persons who teach in the elementary

school from similarly educated persons who do not teach there or who teach in only distantly related areas at much different levels cannot draw confidence as a measure of skill in solving problems in elementary school teaching. Problems designed for a specialized occupation or profession must distinguish members of this group from a random sample of the general college graduate population if these problems are to be considered valid.

One difference between this criterion and the two from which it is derived is that it would require a longitudinal study to demonstrate that persons who graduate from college and enter teaching improve in performance on constructed teaching problems while persons who graduate from college but do not enter teaching remain constant in performance. On the other hand, to show that constructed problems in teaching distinguish between elementary teachers and non-elementary teachers requires only cross sectional studies. The cross sectional approach is the weaker of the two approaches because it is possible to show only that teachers differ in performance from non-teachers, and not that they differ because of teaching experience. To show differences between teachers and non-teachers is crucial, but to show that these differences are at least partly attributable to teaching experience would increase confidence in the validity of the measure employed.

A fourth criterion which may be employed to increase confidence in constructed teaching problems is the ability of a measure of this kind to distinguish particular groups of teachers from the general teaching population. One of these groups consists of demonstration teachers, consulting teachers, and supervising teachers who have been selected for these capacities because of widely recognized skill in teaching. The reason why this group would be expected to perform better than a random sample of teachers in the field is not that their skills are more highly developed than those of other teachers, but rather that the processes used to select this group may be expected to eliminate teachers who perform poorly or whose performance is wholly tied to a particular kind of teaching situation.

A second group of teachers which constructed problems in teaching would be expected to distinguish from other groups includes those who have produced unusually good gains in pupil achievement over a period of years in more than a single teaching situation. Whereas it is neither possible nor desirable to conduct an investigation in which teachers are actually moved from situation to situation, it may be possible to find teachers who have moved to several different situations (quite possibly to obtain better salaries and greater autonomy) and for whom data on the results of their teaching are available. If such a group can be found, then constructed problems in teaching should differentiate between those in the group who have consistently produced good results and those who have not.

The amount of confidence that can be placed in the validity of a test which differentiates the groups suggested above depends almost wholly on the particular procedures

utilized by the investigator. If the procedures used to select the groups are erratic, illogical, or vaguely defined, then comparatively little confidence may be placed in the results. The investigator who undertakes to use the fourth criteria mentioned must be willing to face and resolve a great many problems in developing adequate procedures. For instance, to differentiate those demonstration and supervising teachers who have been selected because they are widely recognized as skillful teachers from those who attain this status for other reasons confronts the investigator with an immense procedural problem, the resolution of which only begins with the examination of the employment practices of the various institutions where such teachers may be found.

A fifth criterion involves the relationships between constructed problems in teaching and other measures of behavior. For instance, if performance on these problems correlates .90 with performance on tests of reading comprehension, one is immediately suspicious that the problems are not good measures of skill in solving teaching problems. It is difficult to think of reasonable explanations for such a high relationship. This does not mean that such relationships might not in fact exist, but rather that one has not provided for their occurrence in the rationale which underlies the test; for this reason confidence in both the test and the rationale which underlies it is seriously weakened.

The kind of criterion just mentioned is closely linked to an additional set of criteria which make up the basis for the construct validity of a test (12). The authors do not wish to engage in a general discussion and evaluation of construct validity or the criticisms of this type of validity (3). They seek only to show how construct validity may be employed in a limited way to increase confidence in the validity of constructed problems in teaching.

In the earlier pages of this paper the authors set forth a number of assumptions and hypotheses centering around the notion that good problem solving in teaching requires that the teacher acquire many instrumental responses and that he be able to choose the most appropriate of these responses on the basis of a careful utilization of the information available in the particular situation in which the response is to be made. These assumptions form a sizable portion of rationale upon which the notion of constructed problems in teaching was originated by the authors, and they form a particular way of looking at what teachers do and what it is that leads to success in teaching. What the authors have said is that skill in teaching depends on the instrumental responses that teachers acquire and on how well teachers utilize the information that is available to them. They have also said that the acquisition of instrumental responses useful in solving a wide scope of teaching problems depends on such factors as learning ability, autonomy in the early years of teaching, breadth of goals learned in preparatory institutions, and so on. If what they have said is true, then they should be able to show that it is true by means of the constructed teaching problems they

design. This seems straight-forward and clear enough, but actually an important problem is buried in it. The authors hope to show in a few sentences what this problem is.

If one sets up a series of assumptions and hypotheses about a teacher's classroom behavior and then constructs a test (such as constructed problems in teaching) to show that these assumptions and hypotheses are true, he can demonstrate that they are true of test performance without ever actually demonstrating that they are true of the classroom behavior. In other terms, an investigtor can attempt to show his test is valid by showing that the results of the test verify hypotheses he has made about the behavior he is measuring. But it is clear that to attempt to validate a test by this means is essentially to engage in logical circularity. At some point one must break out of the circle.

The first of two means the authors suggest for avoiding circularity is to utilize some criteria for the validity of constructed problems in teaching which are independent of the rationale upon which the test is based. For instance, the criterion of differentiating teachers from non-teachers and the criterion that demonstration and supervising teachers will out-perform random samples of teachers are independent of the assumptions and hypotheses made by the authors. Any measure of skill in teaching would be expected to meet these criteria, regardless of its rationale.

The second means of avoiding circularity is to utilize problems which have a very high face validity for teaching. This is to say, constructed problems in teaching must as nearly as possible replicate the problem situations which a teacher might encounter in the classroom. The extreme importance of high face validity must not be underestimated. The problem of circularity is generated because of the discrepancy between test behavior and actual (or criterion) behavior, and the closer the test problems come to the actual problems with which teachers must deal, the less discrepancy there is between test behavior and actual classroom behavior. The less the discrepancy, the greater the probability that the responses elicited by the test problems are those also elicited by actual classroom problems.

As an instance of what the authors mean by high face validity, a specific example may be used. Elementary teachers are regularly confronted with having to diagnose pupil errors in reading by listening to the pupil read orally. The initial phase of the problem is for the teacher to identify and classify such errors accurately. A problem of this kind can be replicated outside the classroom by tape recording the oral reading of the pupil and having the teacher listen to the recording and then identify and classify the errors. Such a problem has extremely high face validity, although it is not identical to the diagnostic problem which a teacher would encounter in the classroom, since some visual cues are missing. A researcher who wants to provide such cues can make a motion picture of the child reading to accompany the sound and thereby provide still greater face validity.

The possibility of engaging in circularity in trying to increase the confidence in the validity of constructed teaching tasks, then, can be avoided by utilizing criteria independent of those which are stated in the rationale underlying the problems and by using problems which have high face validity. Once one is able to show that constructed teaching problems are valid with reference to meeting certain empirical criteria, i.e., the differentiation of teachers from non-teachers, the differentiation of demonstration and supervising teachers from other teachers, and the improvement of performance with teaching experience and failure to improve without it, he is able to increase confidence in the validity of the test problems by showing that the test confirms those assumptions and hypotheses which make up the content of construct, i.e., skill in solving teaching problems, upon which he has originally predicated the test.

To restate the matter, the authors have said that skill in solving teaching problems depends on the acquisition of instrumental responses and upon ability to utilize information to select the appropriate instrumental responses in a given situation. They have further said that the acquisition of instrumental responses depends on learning ability, autonomy in the early years of teaching, breadth of goals learned in preparatory institutions, etc. They have then gone ahead and constructed problems by which to measure skill in solving teaching problems. If these constructed problems measure what the authors say they will measure, they will confirm the assumptions and hypotheses the authors have made and will increase confidence in the validity of the measure. If the constructed problems do not confirm the assumptions and hypotheses, they will decrease confidence in the validity of the measure.

It should be noted at this point that the authors are not using construct validity as the basic means of test validation, because to do so probably involves circularity. Rather, they are using construct validity as a means of increasing confidence in the validity of a test which has been partially validated against empirical criteria. Ultimately, full confidence in the validity of constructed teaching problems can be obtained by demonstrating that an increase in performance on constructed teaching problems leads to increased skill in solving a wider scope of problems in the classroom.

The objective of the strategy which the authors have proposed is to answer the question concerning how well a teacher solves those problems which characterize elementary school teaching as a professional endeavor but do not equally well represent all jobs in elementary school teaching. The authors have attempted to show that in order to attain this objective certain assumptions and hypotheses about teaching and teachers need to be made. These assumptions and hypotheses center around the notion that teacher behavior is problem-solving behavior and that teaching means solving problems. Once it is assumed that teacher behavior is problem-solving behavior and that teaching means solving problems, it follows that in order to assess how well teachers perform it is necessary to confront them with teaching problems to solve, and

that in order to compare how well teachers perform in relation to each other, which is implicit in the objective assumed, it is necessary that these problems be the same for every teacher.

To get problems which are the same for every teacher, it is necessary to construct them. But in the process of constructing problems it is necessary to assume a limited and consistent set of educational objectives to serve as a basis for generating the problems. Because many different kinds of educational objectives are assumed by teachers, not all of them consistent with one another and some more widely assumed than others, it becomes impossible to represent all possible jobs in teaching. This fact precludes the possibility of predicting to all teaching jobs on the basis of performance of constructed teaching problems, and this means that the objectives toward which these problems are directed must differ from that objective assumed by the traditional strategies in teacher effectiveness research, which is, at least ideally, to match teachers with the full range of particular jobs.

When the researcher uses constructed problems in teaching, he also becomes charged with the responsibility of showing that the problems he constructs are valid teaching problems and that they may legitimately be used for the purposes for which he wishes to use them, namely, to identify those teachers who can best solve the widest scope of problems in teaching, that is, those who show the greatest skill in the practice of their profession. The problem of showing that constructed problems in teaching are valid is a complex one. It is complex because there are few criteria which can both inspire confidence and be defined by those procedures available to the investigator at the present time. The problem is not one of being unable to think of criteria; this is not difficult. The problem is to be able to define by the procedures used exactly what the criterion is without destroying confidence in it. Those criteria which the authors believe can be operationally defined and can draw confidence have been suggested in the foregoing section and are those which the authors utilize in the validation of constructed problems in teaching.

QUESTION 4: How Can One Predict How Well a Particular Elementary School Teacher Will Solve Those Problems Which Characterize Elementary School Teaching as a Professional Endeavor But Which Do Not Equally Well Represent All Jobs in Elementary School Teaching?

The fourth question with which persons charged with selecting and educating prospective teachers are centrally concerned is how one can predict how well a particular elementary school teacher will solve those problems which characterize elementary school teaching as a professional endeavor but which do not equally well represent all jobs in elementary school teaching. In the estimation of the authors, the possi-

bility of answering this question rests almost wholly on whether constructed problems in teaching can be shown to successfully sort out from the general population of teachers that particular group which best performs the widest scope of problems in teaching. If it can, then at least two, and possibly three, prediction procedures may be utilized.

The first problem confronted by a preparatory institution is to select students who are most likely to succeed in subsequent teaching positions. The staff of the preparatory institution does not know the full range of particular teaching jobs in which its graduates will be placed; therefore, it may not select students who can succeed only under this or that particular set of conditions. The matter is, of course, not only one of subsequent placement, but also one of selecting students who can adapt to changing teaching practices and changing educational goals as they mature as teachers in a changing society.

In selecting potential teachers, the best policy, from the viewpoint of predicting subsequent success, would be to select those students who possess the characteristics of teachers who have demonstrated the greatest skill in solving teaching problems. What these characteristics are depends on what characteristics good solvers of teaching problems reveal. The authors have hypothesized that one of the characteristics is a relatively high degree of learning ability, the ability to acquire instrumental responses; a second characteristic is finding high reward value in the perception of changes in pupil behavior. Aside from these characteristics the identification of the characteristics of good solvers of teaching problems is a matter of discovery and can be dealt with in part by utilizing measures now available and in part by creating new measures suggested by the factors that appear to enter into problem-solving performance as this performance is observed under laboratory conditions.

Once students have been selected and have entered the teacher-preparation curriculum, the basis of prediction may be shifted from the characteristics of good problem solvers to performance on constructed teaching problems. Prediction on the basis of problem-solving performance may be done in one of two ways, depending on the information about problem-solving performance accumulated at the time the prediction is to be made.

When only limited information bearing on the relationship between problem-solving performance as an undergraduate and problem-solving performance as a teacher is available, "blanket" predictions (i.e., predictions which do not take into account individual differences in rates of growth) can be made. Such predictions can be expected to have modest accuracy, since no attempt is made to adjust predictions according to the pattern of growth in solving teaching problems. When a great deal of information bearing on the relationship between problem-solving performance as an undergraduate and problem-solving performance as a teacher is available, differential prediction (i.e., prediction according to the

sub-class of teachers with respect to pattern of growth in teaching skill) is possible. Such predictions can be expected to be reasonably accurate, since individual differences in patterns of growth in teaching skill can be taken into account.

Whether differential predictions can be made depends, of course, on whether there are different patterns of growth among teachers with respect to the acquisition of teaching skills. The authors hypothesize that there are such patterns. Yet the fact that they make this hypothesis is perhaps less significant than the fact that they can make it with some hope of verifying it--an alternative open to no other current strategy, since one must have a valid, controllable body of teaching problems by which teaching skill can be continuously assessed in order to verify this hypothesis.

REFERENCES IN THE TEXT

1. Barr, A. S.; Eustice, D. E.; and Noe, E. J., "The Measurement and Prediction of Teacher Efficiency," Review of Educational Research 25:261-269, June, 1955.

2. Barr, A. S., and Jones, R. E., "The Measurement and Prediction of Teacher Efficiency," Review of Educational Research 28:256-264, June, 1958.

3. Bechtoldt, Harold P., "Construct Validity: A Critique," The American Psychologist 14:619-630, October, 1959.

4. Bindra, Dalbir, Motivation, A Systematic Reinterpretation, Ronald Press, New York, 1959.

5. Brownell, William A., "The Evaluation of Learning in Arithmetic," Arithmetic in General Education, pp. 231-232, Sixteenth Yearbook of the National Council of Teachers of Mathematics, Bureau of Publications, Teachers College, Columbia University, New York, 1941.

6. Bruner, J. S.; Goodnow, J. J.; and Austin, George A., A Study of Thinking, Wiley and Sons, New York, 1956.

7. Bryan, Glenn L., Electronics Troubleshooting: A Behavioral Analysis, Project Designation NR 153-093, Contract Nonr-228(02), Department of Psychology, University of Southern California, Los Angeles, March, 1956.

8. Burton, W. H., and Brueckner, L. J., Supervision, A Social Process, Appleton-Century-Crafts, New York, 1955.

9. Castetter, D. D.; Standlee, L. S.; and Fattu, N. A., Teacher Effectiveness: An Annotated Bibliography, Bulletin of the Institute of Educational Research, vol. 1, no. 1, Indiana University, Bloomington, 1954.

10. Cohen, Morris R., and Nagle, Ernest, An Introduction to Logic and Scientific Method, Harcourt, Brace, and Co., New York, 1934.

11. Cook, W. W.; Leeds, C. H.; and Callis, R., Minnesota Teacher Attitude Inventory Manual, Psychological Corporation, New York, undated.

12. Cronbach, L. J., and Meehl, P. E., "Construct Validity in Psychological Tests," Psychological Bulletin 52:281-302, July, 1955.

13. Domas, S. J., and Tiedeman, D. V., "Teacher Competence: An Annotated Bibliography," Journal of Experimental Education 19:101-218, December, 1950.

14. Greene, Harry A., and Gray, William S., "The Measurement of Understanding in the Language Arts," in The Measurement of Understanding, pp. 175-212, Forty-fifth Yearbook, National Society for the Study of Education, Public School Publishing Co., Bloomington, Ill., 1946.

15. Gulliksen, Harold, "Intrinsic Validity," The American Psychologist 5:511-517, October, 1950.

16. Havighurst, Robert J., Human Development and Education, Longmans, Green and Co., New York, 1953.

17. Kearney, Nolan C., Elementary School Objectives, Russell Sage Foundation, New York, 1953.

18. Medley, D. M., and Mitzel, H. E., "Some Behavioral Correlates of Teacher Effectiveness," Journal of Educational Psychology 50:239-246, December, 1959.

19. Mitzel, Harold E., "Criteria of Teacher Effectiveness," in Encyclopedia of Educational Research, pp. 1481-1485, The Macmillan Co., New York, 1960.

20. Monroe, Walter S., Teaching-Learning Theory and Teacher Education, 1890 to 1950, University of Illinois Press, Urbana, 1952.

21. Morsh, J. E., and Wilder, E. W., Identifying the Effective Instructor: A Review of the Quantitative Studies, 1900-1952, Research Bulletin, Air Force Personnel and Training Research Center, Lackland AFB, San Antonio, Texas, October, 1954.

22. Remmers, H. H., et al., "Report of the Committee on the Criteria of Teacher Effectiveness," Review of Educational Research 22:238-263, June, 1952.

23. Remmers, H. H., et al., "Second Report of the Committee on the Criteria of Teacher Effectiveness," Journal of Educational Research 46:641-658, May, 1953.

24. Ryans, David G., "The Criteria of Teacher Effectiveness," Journal of Educational Research 42:690-699, May, 1949.

25. Ryans, David G., "The Investigation of Teacher Characteristics," Educational Review 34:371-396, October, 1953.

26. Ryans, David G., "Theory Development and the Study of Teacher Behavior," Journal of Educational Psychology 47:462-475, December, 1956.

27. Ryans, David G., "Prediction of Teacher Effectiveness," in Encyclopedia of Educational Research, pp. 1486-1491, The Macmillan Co., New York, 1960.

28. Sears, R. R., "A Theoretical Framework for Personality and Social Behavior," The American Psychologist 6:476-483, September, 1951.

29. Skinner, B. F., Science and Human Behavior, The Macmillan Co., New York, 1953.

30. Smith, B. O., A Study of Teaching, Twelfth Annual Meeting, American Association of Colleges for Teacher Education, Chicago, February, 1960, mim.

31. Standlee, L. S., and Popham, W. J., Preparation and Performance of Teachers, Bulletin of the School of Education, vol. 34, no. 6, Indiana University, Bloomington, November, 1958.

32. Tiedeman, D. V., and Cogan, Morris, "New Horizons in Educational Research," Phi Delta Kappan 39:286-291, March, 1958.

33. Turner, R. L., and Fattu, N. A., Problem Solving Proficiency Among Elementary School Teacher, I, The Development of Criteria, Monograph of the Institute of Educational Research, Indiana University, Bloomington, April, 1960 (In press).

34. Turner, R. L., Problem Solving Proficiency Among Elementary School Teachers, II, Teachers of Arithmetic Grades 3-6, Monograph of the Institute of Educational Research, Indiana University, Bloomington, June, 1960 (In press).

35. Wiles, Kimball, Supervision for Better Schools, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1955.

36. White, Robert W., "Motivation Reconsidered: The Concept of Competence," Psychological Review 60:297-333, September, 1959.

BULLETINS IN THE FIELD OF EDUCATION
INDIANA UNIVERSITY

The Bulletin of the School of Education, Indiana University, was first published in 1924 and has appeared regularly since that time. A complete list of bulletins may be obtained from the School of Education upon request. The studies included in the present and in the two volumes immediately preceding it in date are listed below. Unless otherwise indicated, these may be obtained for $1 each from the Indiana University Bookstore, Bloomington, Indiana.

Volume 34 (1958)

1. Franzén, Carl G. F., Foreign Language in the Curriculum, 30 pp.

2. Mercille, Margaret G., Undergraduate Workshop in Elementary Education, 40 pp.

3. Strom, Ingrid M., The Role of Literature in the Core Curriculum, 60 pp.

4. Popham, W. James, and Standlee, Lloyd S., Out-of-School Activities and Professional Performance of Teachers, 38 pp.

5. Fay, Leo C., Improving the Teaching of Reading by Teacher Experimentation, 104 pp.

6. Standlee, Lloyd S., and Popham, W. James, Preparation and Performance of Teachers, 48 pp.

Volume 35 (1959)

1. Wright, Wendell W., and Jung, Christian W., Why Capable High School Students Do Not Continue Their Schooling, 78 pp.

2. Hughes, Otto, The Role of the Campus Laboratory School, 48 pp.

3. Research Studies Related to the Improvement of Education in Thailand, 66 pp.

4. Batchelder, Howard T., Report of a State-Wide Study of the Public Schools and Teacher Education in Indiana, 82 pp.

5. Church, Harold H., and Lewis, Melvin S., An Appraisal of the School Surveys Conducted by the School of Education, Indiana University, 66 pp.

6. Barr, W. Monfort; Church, Harold H.; Stapley, Maurice E.; and McGhehey, Marion A., School District Reorganization in Indiana, 136 pp.

Volume 36 (1960)

1. Clayton, A. Stafford, editor, John Dewey in Perspective, 44 pp., $1.25.

2. Klinge, Paul E., A University Summer Program for Gifted Science Students, 54 pp., $1.25.

3. Turner, Richard L., and Fattu, Nicholas A., Skill in Teaching, A Reappraisal of the Concepts and Strategies in Teacher Effectiveness Research, 39 pp., $1.25.